



①9 BUNDESREPUBLIK  
DEUTSCHLAND



DEUTSCHES  
PATENTAMT

⑫ Patentschrift  
⑩ DE 44 06 020 C 1

⑤1 Int. Cl.<sup>6</sup>:  
**G 06 K 9/62**  
G 07 C 9/00

②1 Aktenzeichen: P 44 06 020.3-53  
②2 Anmeldetag: 24. 2. 94  
④3 Offenlegungstag: —  
④5 Veröffentlichungstag  
der Patenterteilung: 29. 6. 95

DE 44 06 020 C 1

Innerhalb von 3 Monaten nach Veröffentlichung der Erteilung kann Einspruch erhoben werden

⑦3 Patentinhaber:

Zentrum für Neuroinformatik GmbH, 44801 Bochum,  
DE

⑦2 Erfinder:

Konen, Wolfgang, Dr., 44789 Bochum, DE;  
Vorbrüggen, Jan C., 44809 Bochum, DE; Würtz, Rolf  
P., 44799 Bochum, DE

⑤6 Für die Beurteilung der Patentfähigkeit  
in Betracht gezogene Druckschriften:

WO 93 15 475  
BICHSEL, M., SEITZ, P.: DAGM,  
Springer-Verlag 1990, S. 106-113;  
LODES et al.: IEEE Transaction on Computers, 42,  
1993, pp. 300-311;  
BUHMANN et al.: IJCNN II, San Diego 1990,  
pp. 411-416;

⑤4 Verfahren zur automatisierten Erkennung von Objekten

⑤7 Die Erfindung betrifft ein verbessertes bildverarbeitendes Verfahren zur automatisierten Erkennung oder Verifikation von Objekten aus digitalen Bildern dieser Objekte. Dieses Verfahren ist dadurch gekennzeichnet, daß man aus einem digitalen Bild mit Gabor-Filtern verschiedener Größe und Orientierung Gabor-Merkmale extrahiert, welche auf einem verschieb-, skalier- und deformierbaren Gitter, dem Graphen, angeordnet werden, dessen optimale Form durch die zweiphasige Minimierung einer Graphen-Vergleichsfunktion bestimmt wird. In der ersten Phase werden sowohl Größe als auch Position des Graphen simultan optimiert, in der zweiten Phase wird die intrinsische Form des Graphen optimiert. Das Verfahren ist gegenüber dem Stand der Technik dahingehend verbessert, daß gleiche Objekte auch bei unterschiedlicher Abbildungsgröße erkannt werden und daß, besonders für die Aufgabe der Verifikation, ein wesentlich beschleunigtes Optimierungsverfahren angegeben werden konnte. Die Erfindung betrifft ferner den Einsatz automatisierter bildverarbeitender Gesichtserkennungsverfahren zur Anwendung in Personenzugangskontrollsystemen oder anderen Einrichtungen zur Identitätsüberprüfung.

DE 44 06 020 C 1

Die Erfindung betrifft ein automatisiertes bildverarbeitendes Verfahren zur größen- und lageinvarianten Erkennung von intrinsisch deformierbaren Objekten, insbesondere Gesichtern.

Seit einigen Jahren sind auf dem Gebiet der Bildverarbeitung Verfahren bekannt, die eine Erkennung von Objekten aus einzelnen Ansichten (Bildern) dieser Objekte erlauben. Diese Verfahren zeichnen sich durch eine mehr oder weniger starke Robustheit gegenüber Variationen der Objektansichten (Lage, Orientierung im Raum, Größe, objektintrinsische Verzerrungen, Beleuchtung) aus. Insbesondere beschäftigen sich zahlreiche Arbeiten mit dem schwierigen Problem der Erkennung von Gesichtern, welche eine Klasse von Objekten mit zahlreichen intrinsischen Deformationsfreiheitsgraden (Mimik) darstellen.

In den zahlreichen Arbeiten in einschlägigen Fachzeitschriften, die sich mit der Erkennung von Gesichtern beschäftigen und die in einschlägigen Fachzeitschriften publiziert sind, werden folgende Lösungsansätze vorgeschlagen:

1. Bichsel und Seitz, DAGM, S. 106—113, Springer-Verlag 1990 haben vorgeschlagen, aus Bildern extrahierte Merkmale einem Neuronalen Netz (Mehrlagen-Perzeptron) zuzuführen. Ein sehr ähnlicher Ansatz wird in Bouattour, Fogelman-Soulie et al., Artificial Neural Networks 2, S. 456, Elsevier-Verlag 1992 gewählt. In Lampinen, Artificial Neural Networks, S. 328, Elsevier-Verlag 1991 wird zur Merkmalsklassifikation die Faktorenanalyse und ein neuronales Kohonen-Netz benutzt. Der Nachteil all dieser Verfahren ist, daß für jede neu zu erkennende Person eine ganze Reihe von Ansichten aufgenommen werden muß und daß die Neuronalen Netze sodann erneut trainiert werden müssen. Außerdem werden Ergebnisse dieser Verfahren in Form von Erkennungsraten nur für kleine Datenbanken (3—20 Personen) berichtet. Es muß davon ausgegangen werden, daß die Erkennungsrate mit steigender Personenzahl abnimmt.

2. Aus Lades et al., IEEE Transaction on Computers, 42, 300—311, 1993 [Lades'93] ist ein Ansatz bekannt, bei dem der Vergleich zwischen zu erkennendem und gespeichertem Bild durch einen flexiblen Abbildungsmechanismus realisiert wird, wobei die bestmögliche Abbildung durch ein Optimierungsverfahren bestimmt wird. Dieses Verfahren ist sehr leistungsfähig in bezug auf Lage, Orientierung im Raum und objektintrinsische Verzerrungen, ferner werden auch bei großen Personenzahlen (> 100) hohe Erkennungsraten erreicht. Es hat jedoch den Nachteil, daß es keine größeninvariante Erkennung erlaubt.

3. Ein interessantes Verfahren, das ebenfalls eine Abbildung zwischen Bildern durch ein Optimierungsverfahren realisiert, wird von M. Shackelton, Patentschrift WO 93/15475, 1993 [Shackelton'93] beschrieben. Das Verfahren findet in Bildern automatisch hervorspringende Merkmale und verbindet diese zu einer netzartigen Struktur (Template). Solche Templates können prinzipiell dazu benutzt werden, ähnliche Bilder miteinander zu vergleichen und können somit auch zum Vergleich von Gesichtern herangezogen werden. Es werden allerdings in [Shackelton'93] keine weiteren Angaben dazu ge-

macht, wie ein solcher Vergleich verschiedener Bilder gleicher Objekte im Detail aussehen kann und welche Güte der Klassifizierung damit erreicht werden kann.

Bezogen auf Personenzugangskontrollen gibt es im Stand der Technik keine gewerblichen Anwendungen, die die automatisierte, lage- und größeninvariante Erkennung von Gesichtern als Kontrollmerkmal einsetzen.

Zum besseren Verständnis der hier vorliegenden Erfindung wird im folgenden der Stand der Technik, wie er durch [Lades'93] vorgegeben ist, detaillierter dargestellt. Dazu werden zunächst einige für die Erfindung wesentlichen Begriffe und Bezeichnungen zusammengefaßt:

Es ist Stand der Technik, wie er beispielsweise in der Arbeit von D. J. Field, Journal Opt. Soc. Am. A4, S. 2379—2394, 1987 dargestellt ist, daß zur Merkmalsextraktion aus Bildern Filteroperationen (mathematisch: Faltungen) mit bestimmten Filtermasken verwendet werden. Eine bestimmte Klasse von Filtermasken sind die Gabor-Filter. Verschiedene Gabor-Filter unterscheiden sich in Größe und Orientierung in der Bildebene:

— Ein Gabor-Merkmal bezeichnet das Ergebnis einer Faltung des Bildes an einem bestimmten Pixel mit einem bestimmten Gabor-Filter.

— Ferner bezeichnet, wenn ein Bild mit  $k$  verschiedenen Gabor-Filter gefaltet wurde, ein Jet die Gesamtheit aller Gabor-Merkmale an einem bestimmten Pixel, die in dem Jet zu einem  $k$ -komponentigen Vektor zusammengefaßt werden.

— Ein regulärer Graph ist die Zusammenfassung von  $m \times n$  Jets in einer zweidimensionalen, gitterförmigen Anordnung (s. Bild 1a). Jeder Jet ist einem der Knoten des Gitters zugeordnet; benachbarte Knoten werden durch horizontale und vertikale Verbindungen miteinander verknüpft. Für zu speichernde Objektbilder wird die Information über ein Objektbild in Form von regulären Graphen abgelegt.

— Ein irregulärer Graph besteht ebenfalls aus  $m \times n$  Knoten, die in gleicher Topologie wie beim regulären Graphen verknüpft sind, jedoch liegen hier die Knoten nicht notwendigerweise auf einem rechteckigen Gitter, sondern sind beliebig in der Bildebene angeordnet (s. Bild 1b). Es werden den Knoten jeweils die Jets zugeordnet, die zu dem Pixel gehören, auf dem der Knoten zu liegen kommt.

— Die horizontalen und vertikalen Verbindungen zwischen benachbarten Knoten werden auch als Links bezeichnet.

— Zur Bewertung der Ähnlichkeit zweier Graphen (regulär oder irregulär) wird eine Graphen-Vergleichsfunktion  $E$  eingeführt:

$$E = E_{\text{sim}} + \lambda E_{\text{top}} \quad (1)$$

Diese Funktion nimmt um so niedrigere Werte an, je ähnlicher zwei Graphen sind. Der erste Term  $E_{\text{sim}}$  bewertet die Ähnlichkeit der Jets an jeweils korrespondierenden Knoten der beiden Graphen, indem die jeweils korrespondierenden Gabor-Merkmale miteinander verglichen werden (negativer Kosinus des Winkels zwischen beiden Jets); ho-

he Ähnlichkeit bedeutet dabei niedrige Werte von  $E_{\text{sim}}$ . Der zweite Term  $E_{\text{top}}$  bewertet die topologische Ähnlichkeit beider Graphen, indem die Beträge der Differenzvektoren jeweils korrespondierenden Links aufsummiert werden. (Ein anschauliches Modell besteht darin, sich die Links als mechanische Federn vorzustellen, die in ihren Ruhelagen den ersten der beiden Graphen bilden.  $E_{\text{top}}$  ist dann proportional zur Arbeit, die gegen die Federkräfte aufgebracht werden muß, um die Links mit der Topologie des zweiten Graphen zur Deckung zu bringen.)

— Werden beim Vergleich eines neuen Objektbildes mit  $N$  gespeicherten Objekten  $N$  Werte der Graphen-Vergleichsfunktion  $E_1 < E_2 < \dots < E_N$  ermittelt, so bezeichnen  $\bar{E}$  und  $\sigma_E$  den Mittelwert und die Standardabweichung der Daten ( $E_2, \dots, E_N$ ), und Objekt 1 ist im Sinne der Vergleichsfunktion das ähnlichste Objekt. Die Ähnlichkeit wird als signifikant bezeichnet und wir sprechen von signifikanter Erkennung, wenn für fest gewählte Parameter  $s_1, s_2$  mindestens eines der Kriterien  $(E_1 - \bar{E})/\sigma_E > s_1$  oder  $(E_1 - E_2)/\sigma_E > s_2$  erfüllt ist. Andernfalls sprechen wir von insignifikanter Erkennung, und das neue Objekt wird als unähnlich zu allen gespeicherten Objekten zurückgewiesen.

In [Lades'93] wird nun ein Verfahren beschrieben, das Gabor-Merkmale verschiedener Orientierung und Skalierung aus einem Bild extrahiert und diese für jedes Pixel in Form von Jets zusammenfaßt. Für ein zu speicherndes Objektbild werden die Jets in einem regulären Graphen, der auf das Objekt zentriert ist, zusammengefaßt. In einem neuen Bild wird ein optimal zu dem gespeicherten Graphen passender Graph gesucht. Dabei enthält der neue Graph an seinen Knoten jeweils die Jets, die zu dem Pixel am Ort des Knoten gehören. Seine Position und Form wird in einem Zweiphasenverfahren optimiert:

1. Verschiebung des Graphen als Ganzes (Global Move), bis ein Optimalzustand gefunden wird.
2. Ausgehend vom Optimalzustand aus 1. werden einzelne Knoten des Graphen verschoben (lokale Verzerrung).

In beiden Phasen läuft die Optimierung folgendermaßen ab: Es wird ein zufälliger Verschiebungsvektor ausgewählt und auf den Graphen bzw. einen Knoten des Graphen angewendet. Der Zustand des so veränderten Graphen wird mittels der Graphen-Vergleichsfunktion  $E$  bewertet. Nur wenn der neue Zustand einen günstigeren Wert für  $E$  hat, wird die Verschiebung akzeptiert.

Beide Phasen terminieren, wenn in  $V_{\text{max}}$  aufeinanderfolgenden Verschiebungsschritten keine Verbesserung gefunden wird. Das Ergebnis der Optimierung ist der finale Wert der Graphen-Vergleichsfunktion  $E$ . Die Optimierung wird für alle gespeicherten Graphen wiederholt, und man erhält eine Sequenz von Werten  $E_1 < E_2 < E_3 < \dots$  anhand welcher bestimmt werden kann, ob eine signifikante Erkennung von Objekt 1 vorliegt.

Das beschriebene Verfahren hat jedoch folgende Nachteile:

- Es ist nicht größeninvariant, da der Graph nur verschoben, nicht skaliert wird. Es wurde bisher in der Literatur (Buhmann et al., IJCNN II, 411–416, IEEE, San Diego, 1990) davon ausgegangen, daß

bei Skalierung des Graphen auch die Filterkoeffizienten entsprechend inter- bzw. extrapoliert werden müssen. Ein solches Verfahren ist rechenaufwendig, da nach jedem Optimierungsschritt interpoliert werden muß. Ferner wurde bisher davon ausgegangen, daß die Größenbestimmung nur durch ein hierarchisches Verfahren, das von groben zu feinen Auflösungsstufen fortschreitet, gelingt. Die hier beschriebene Erfindung löst die Aufgabe der größeninvarianten Erkennung durch ein deutlich vereinfachtes Verfahren.

— Der Suchaufwand wächst linear mit der Anzahl der gespeicherten Objekte. Ferner wird die korrekte Erkennung mit zunehmender Objektanzahl immer schwieriger. Damit wird der Einsatz in großen Datenbanken problematisch. Die hier dargestellte Erfindung beschreibt ein neues Verfahren, die Verifikation, welches die obigen Nachteile vermeidet. Dieses Verfahren läßt sich darüber hinaus besonders gut mit der größeninvarianten Erkennung kombinieren.

Die aus dem Stand der Technik bereits bekannten Verfahren weisen eine Reihe von Nachteilen auf und vermögen nicht, in jeder Hinsicht zu befriedigen. Es besteht daher ein ständiger Bedarf an verbesserten Verfahren zur Erkennung intrinsisch deformierbarer Objekte.

Aufgabe der vorliegenden Erfindung war es, ein gegenüber dem bekannten Stand der Technik verbessertes Verfahren zur Erkennung von Objekten zur Verfügung zu stellen. Dabei war angestrebt, eine Möglichkeit zur Differenzierung dahingehend zu schaffen, daß verschiedene Bilder desselben Objektes (mit Unterschieden in Lage, Ansicht und Größe des Objektes) als ähnlich erkannt werden, während Bilder von verschiedenen Objekten als unähnlich klassifiziert werden.

Eine weitere Zielsetzung bestand darin, eine technische Verwendung dieses Verfahrens als Personenzugangskontrollsystem zu entwickeln, das auf einem verbesserten Verfahren zur Erkennung von Gesichtern basiert. Dabei besteht die Aufgabe insbesondere darin, die Überprüfung der Zugangsberechtigung auch bei einem großen Kreis berechtigter Personen schnell und sicher durchführen zu können.

Überraschenderweise wurde gefunden, daß man die größeninvariante Objekterkennung einfach dadurch deutlich verbessern kann, daß man den Global Move simultan mit einem Global Scale verbindet, indem man den Graphen in jedem Optimierungsschritt sowohl verschiebt als auch mit einem Faktor nahe 1 skaliert. Eine Interpolation der Filterkoeffizienten ist entgegen der Erwartung für einen weiten Bereich von Objektgrößen (von etwa 60% bis 140% der Originalgröße) nicht notwendig. Damit hat man eine gegenüber bestehenden Lösungen beschleunigte Anpassung auf Größenänderungen, wie sie für viele Anwendungen von besonderer Bedeutung sind.

Gegenstand der vorliegenden Erfindung ist ein Verfahren zur automatisierten Erkennung von Objekten aus Bildern dieser Objekte, wobei man aus einem digitalen Bild mit Gabor-Filtern verschiedener Größe und Orientierung Gabor-Merkmale extrahiert, die in einem verschieb-, skalier- und deformierbarem Graphen  $G$  zusammengefaßt werden. Für jeden der gespeicherten Graphen optimiert man die Form und Lage des Graphen  $G$  dahingehend, daß die Graphen-Vergleichsfunktion  $E$  optimale Werte annimmt. Die Auffindung des

optimalen Graphen G führt man dabei für jeden der gespeicherten Graphen in zwei Phasen durch.

#### Phase 1

Auswahl eines zufälligen Verschiebungsvektors und Skalierungsfaktors für den gesamten Graphen. Der so veränderte Graph im neuen Bild wird über die Graphen-Vergleichsfunktion E mit dem gespeicherten Graphen verglichen, wobei die Links des gespeicherten Graphen zuvor mit dem Skalierungsfaktor multipliziert werden. Nur wenn der neue Zustand einen günstigeren Wert für E hat, wird die Veränderung des Graphen akzeptiert.

#### Phase 2

Alle Links des gespeicherten Graphen werden mit dem in Phase 1 ermittelten optimalen Skalierungsfaktor multipliziert. Ausgehend vom in Phase 1 ermittelten Optimalzustand des neuen Graphen werden einzelne Knoten dieses Graphen verschoben (lokale Verzerrung), bis ein optimaler irregulärer Graph gefunden wird (Bild 2).

In einer bevorzugten Ausführungsform der vorliegenden Erfindung wird in jedem Optimierungsschritt der Phase 1, ausgehend von der momentanen Graphen-Lage und -Größe, zufällig eine Verschiebung von bis zu 3% der Bildgröße und gleichzeitig eine Skalierung des Graphen von bis zu 10% der Graphengröße gewählt. Allerdings liefert das Verfahren bei den untersuchten Bildern in einem weiten Bereich von Parametern annähernd gleich gute Ergebnisse.

Mit dieser Ausführungsform werden für Bilder, die gespeicherte Objekte in einer um 75% kleineren Skalierung zeigen, die richtigen Skalierungsfaktoren mit einer Genauigkeit von  $\pm 2\%$  ermittelt. Dies zeigten Versuche, in denen man 12 Gabor-Filter (4 äquidistante Orientierungen mit je 3 verschiedenen Größen, gekennzeichnet durch ihre Optimalfrequenzen  $f$ , die, beginnend bei der höchsten Frequenz von  $f = \pi/2$ , im Abstand von halben Oktaven aufeinanderfolgen) verwendet. Die 12-komponentigen Jets werden in einem  $7 \times 10$ -Graphen angeordnet und als Bildmaterial werden Gesichtsbilder ( $128 \times 128$  Pixel) von Personen vor weißem Hintergrund verwendet, wie sie in [Lades'93] dargestellt sind. Unter diesen Voraussetzungen wird von verschiedenen Anfangslagen und -größen des Graphen G aus die optimale Graphengröße (also hier um 75% kleiner) mit einer Reproduzierbarkeit von  $\pm 2\%$  angefahren.

In einer weiteren Ausführungsform der Erfindung wird die Optimierung des Graphen G nur einmal in Hinblick auf einen der gespeicherten Graphen (Referenzgraph) durchgeführt und bleibt dabei auf die Phase 1 beschränkt. Der so ermittelte optimale Graph G wird dann ohne weitere Optimierungsschritte mit allen gespeicherten Graphen verglichen. Überraschenderweise wurde gefunden, daß trotz des drastisch reduzierten Optimierungsaufwand fast gleichbleibend gute Erkennungsleistungen erzielt werden. Diese Ausführungsform hat den Vorteil, daß sie bedeutend schneller durchgeführt werden kann und damit auch in der Anwendung auf große Datenbanken gespeicherter Graphen geeignet ist.

Neben der Erkennung (Suchen eines Objektes unter einer Anzahl gespeicherter Objektbilder) eignet sich das Verfahren auch besonders gut zur Verifikation, d. h. zur Entscheidung, ob ein vorgegebenes Objektbild mit einem bestimmten gespeicherten Objektbild B (dessen

Graph als Referenzgraph benutzt wird) übereinstimmt. Dazu wird das vorgegebene Bild O außer mit dem gespeicherten Bild B auch mit N zufällig ausgewählten anderen gespeicherten Bildern verglichen, um zu entscheiden, ob die Ähnlichkeit zwischen B und O signifikant ist. Dabei ist N eine feste, nicht von der Gesamtgröße der Datenbank abhängige Zahl, so daß in diesem Verfahren der Rechenaufwand unabhängig von der Größe der Datenbank ist. Der Vergleich mit N weiteren gespeicherten Bildern bietet den Vorteil, daß Unterschiede, die durch eine geänderte Aufnahmesituation beim vorgegebenen Bild (z. B. durch andere Beleuchtung) entstehen, getrennt werden von Unterschieden in den Objekten selbst.

Gerade für diese Verifikationsaufgabe eignet sich die oben beschriebene Ausführungsform, bei der zur Reduktion des Rechenaufwandes die Optimierung nur mit einem Referenzgraphen (nämlich B) durchgeführt und bereits nach Phase 1 abgebrochen wird, besonders gut.

Dies zeigen die folgenden Versuchsergebnisse: 88 Objektbilder von Gesichtern mit ca.  $20^\circ$  zur Seite gedrehtem Kopf werden mit 100 gespeicherten Bildern von geradeaus blickenden Personen verglichen. In einer Ausführungsform der Erfindung werden 40 Gabor-Filter (8 äquidistante Orientierungen mit je 5 verschiedenen Größen, gekennzeichnet durch ihre Optimalfrequenzen  $f$ , die, beginnend bei der höchsten Frequenz von  $f = \pi/2$ , im Abstand von halben Oktaven aufeinanderfolgen) und  $7 \times 10$ -Graphen verwendet. Die Schwellen  $s_1$  und  $s_2$  werden so eingestellt, daß Objektbilder zu 100% zurückgewiesen werden, wenn das zugehörige gespeicherte Bild (temporär) aus der Datenbank entfernt ist. Wird nun das zugehörige gespeicherte Bild wieder eingefügt, so ergeben sich mit den gleichen Schwellen folgende Raten der korrekten und signifikanten Erkennung ( $N = 100$ ):

Erkennung, mit Phase 1 und 2	84%
Verifikation, mit Phase 1 und 2	93%
Verifikation, nur Phase 1	91%

Bei einem Versuch zur größeninvarianten Verifikation, in dem 88 Objektbilder mit  $20^\circ$  Kopfdrehung durch 89 Bilder von Gesichtern in reduzierter Größe (75%) ersetzt werden, ergibt sich eine Verifikationsrate (nur Phase 1) von 83%.

Insgesamt erlaubt das Verfahren somit eine automatisierte Erkennung oder Verifikation von Objekten aus digitalen Bildern dieser Objekte unter Benutzung von Gabor-Merkmalen, welche auf einem verschieb-, skalier- und deformierbarem Gitter, dem Graphen, angeordnet sind, dessen optimale Form durch die zweiphasige Minimierung der Graphen-Vergleichsfunktion bestimmt wird. In der ersten Phase werden sowohl Größe als auch Position des Graphen simultan optimiert.

Für viele Anwendungen der visuellen Objekterkennung spielt die Invarianz gegen Größen- und Positionsänderungen von Objekten eine besonderer Rolle. Beispielsweise sind Personen-Zugangskontrollen mit automatischer Kamera ("elektronischer Pförtner") nur dann in der Praxis einsetzbar, wenn Abstand und Position der Person vor der Kamera in bestimmten Grenzen variabel sind. Das hier beschriebene Verfahren eignet sich besonders gut für eine solche Anwendung, zumal auch eine Unterscheidung bei einer großen Zahl von Personen möglich ist. Gegenüber einem menschlichen Pförtner bietet das automatisierte Verfahren die Vorteile

größerer Objektivität, Reproduzierbarkeit und Ermüdungsfreiheit.

Für Personen-Zugangskontrollen müssen hohe Sicherheitsstandards eingehalten werden. Das hier beschriebene Verfahren bietet den Vorteil einer hohen Sicherheit bei der Zurück-Weisung unberechtigter Personen (100% auf dem Testdatenmaterial). Das hier beschriebene Verfahren eignet sich auch dann besonders gut, wenn es in Verbindung mit anderen Zugangskontroll-Maßnahmen (z. B. Codekarten) eingesetzt wird. Durch diese Kombination entstehen folgende Vorteile:

- Eine Person, die unberechtigt in den Besitz einer Codekarte gelangt oder eine solche fälscht, kann alleine durch die Codekarte noch keinen Zugang erlangen. Dies führt zu einer Erhöhung der Sicherheit.
- Durch die Codekarteninformation kann vorgegeben werden, welches der gespeicherten Gesichter zu erkennen ist. Die automatische Erkennung wird reduziert auf die Aufgabe der automatischen Verifikation, ob das aktuelle Bild eine signifikante Ähnlichkeit mit dem gespeicherten Bild aufweist. Im Gegensatz zur Erkennung ist der Rechenaufwand für die Verifikation unabhängig von der Größe der Datenbank; das Verfahren kann somit auch für sehr große Datenbanken eingesetzt werden.

Personen-Zugangskontrollen sind nur ein Beispiel für die technische Anwendbarkeit des Gesichtserkennungsverfahrens. Allgemeiner kann ein Verfahren, bei dem mit einer Kamera ein Gesichtsbild aufgenommen und mit gespeicherten Bildern verglichen wird, zur berührungslosen Identitätsüberprüfung einsetzen. Diese kann dazu benutzt werden, personenbezogene Einstellungen an einem technischen Gerät vorzunehmen (beispielsweise Fahrersitz-Einstellung im Auto) oder personenbezogene Zugangsrechte an einem technischen Gerät einzuräumen.

Besonders vorteilhaft läßt sich ein solches berührungsloses Kontrollverfahren in Kombination mit anderen Identifikationsmerkmalen zur Verifikation einsetzen. Es gelten sinngemäß die gleichen Vorteile wie bei der Verifikation in Personen-Zugangskontrollen.

Das beschriebene Verfahren läßt sich auch vorteilhaft für die Bestimmung der Ähnlichkeit zwischen Objektbildern heranziehen. Als Ähnlichkeitsmaß dient der am Ende der Optimierung erhaltene Wert der Graphen-Vergleichsfunktion. Damit lassen sich große Datenbanken automatisch nach ähnlichen Objekten durchsuchen. Da nur diese für eine weitere (menschliche) Begutachtung herangezogen werden müssen, bedeutet dies einen erheblichen Zeitgewinn. Besonders geeignet ist das Verfahren auch dazu, die Ähnlichkeit zwischen Phantombildern von Gesichtern und gespeicherten Gesichtsfotografien zu ermitteln.

Schließlich läßt sich das Verfahren auch zur automatisierten visuellen Qualitätskontrolle einsetzen. In Produktionsprozessen steht man oft vor der Aufgabe, daß die Übereinstimmung aktuell vorliegender Teile mit den Sollangaben abgeprüft werden muß (Beispiel Bestückung elektronischer Platinen). Oft scheitert die Automatisierung der visuellen Prüfung an der mangelnden Robustheit der eingesetzten Verfahren. Das hier beschriebene Verfahren läßt sich besonders dann vorteilhaft einsetzen, wenn in ihrer grundsätzlichen Struktur ähnliche Produktionsteile sicher auseinandergehalten werden müssen.

1. Verfahren zur automatisierten Erkennung von Objekten aus Bildern dieser Objekte, wobei man aus einem digitalen Bild mit Gabor-Filtern verschiedener Größe und Orientierung Gabor-Merkmale extrahiert, die in einem verschieb-, skalier- und deformierbarem Graphen G zusammengefaßt sind und mit gespeicherten Graphen von anderen Bildern verglichen werden, indem eine Graphen-Vergleichsfunktion E, bestehend aus einem die Ähnlichkeit der Gabor-Merkmale bewertenden Anteil und einem die Formerhaltung des Graphen bewertenden Anteil, berechnet wird, **dadurch gekennzeichnet**, daß man durch schrittweise und simultane Veränderungen von Lage und Größe des Graphen G diesen im Vergleich zu einem gespeicherten Referenzgraph optimal einstellt (Phase 1) und mit diesem G die Werte von E für jeden der gespeicherten Graphen ermittelt, wobei der Abstand des besten Wertes von den übrigen E-Werten eine Entscheidung darüber erlaubt, ob eine signifikante Erkennung vorliegt.
2. Verfahren nach Anspruch 1, wobei man die Optimierung des Graphen G in Phase 1 dahingehend erweitert, daß für jeden der gespeicherten Graphen schrittweise Formveränderungen (lokale Verzerren), die zu einer Verbesserung der Graphen-Vergleichsfunktion E führen, angewandt werden (Phase 2).
3. Verfahren nach Anspruch 1 oder 2, wobei man in jedem Optimierungsschritt der Phase 1 zufällig eine Verschiebung von bis zu 3% der Bildgröße und eine Skalierung des Graphen von bis zu 10% einstellt.
4. Verfahren nach einem der Ansprüche 1 bis 3, wobei man die Größen der eingesetzten Gabor-Filter so wählt, daß ihre jeweiligen Bestfrequenzen im Abstand von Halboktaven aufeinanderfolgen und mindestens drei verschiedene Größen verwendet.
5. Gerät zur berührungslosen Identitätsüberprüfung, bestehend aus Kamera und Computer, dadurch gekennzeichnet, daß ein durch die Kamera aufgenommenes Gesichtsbild durch eines der in den Ansprüchen 1 bis 4 genannten Verfahren mit allen gespeicherten Gesichtsbildern von Personen automatisch verglichen wird, und bei einer signifikanten Erkennung einer der gespeicherten Personen eine personenbezogene Aktion ausgelöst wird.
6. Personen-Zugangskontrolle, bestehend aus Kamera, Computer und elektrischer Zugangsentriegelung, dadurch gekennzeichnet, daß ein durch die Kamera aufgenommenes Gesichtsbild durch eines der in den Ansprüchen 1 bis 4 genannten Verfahren mit allen gespeicherten Gesichtsbildern von zugangsberechtigten Personen automatisch verglichen wird, und bei einer signifikanten Erkennung eines der gespeicherten Gesichter die Zugangsentriegelung betätigt wird.
7. Personen-Zugangskontrolle, bestehend aus Kamera, Computer, elektrischer Zugangsentriegelung und einem davon unabhängigen weiteren Kontrollgerät, durch das die Identität der zu erkennenden Person vorgegeben wird, dadurch gekennzeichnet, daß diese Identität verifiziert wird, indem das zugehörige gespeicherte Gesichtsbild B mit einem durch die Kamera aufgenommenen aktuellen Gesichtsbild nach einem der in den Ansprüchen 1 bis 4

genannten Verfahren verglichen wird und bei positiver Signifikanz die Zugangsentriegelung betätigt wird.

8. Verfahren zur Datenbanksuche in Bilddatenbanken, dadurch gekennzeichnet, daß zu einem vorgelegten Bild nach einem der in den Ansprüchen 1 bis 4 beschriebenen Verfahren alle hinreichend ähnlichen Bilder in der Datenbank gefunden werden, welche sich dadurch auszeichnen, daß für sie die Graphen-Vergleichsfunktion niedrigere Werte als für alle anderen Bilder annimmt.

9. Automatisierte visuelle Qualitätskontrolle im Produktionsprozeß, bestehend aus Kamera, Computer und Aussonderungsmechanismus, dadurch gekennzeichnet, daß die aktuell produzierten Teile über eines der in den Ansprüchen 1 bis 4 beschriebenen Verfahren mit Sollangaben verglichen werden und bei Vorliegen einer Diskrepanz der Aussonderungsmechanismus betätigt wird.

---

Hierzu 1 Seite(n) Zeichnungen

---

20

25

30

35

40

45

50

55

60

65

- Leerseite -

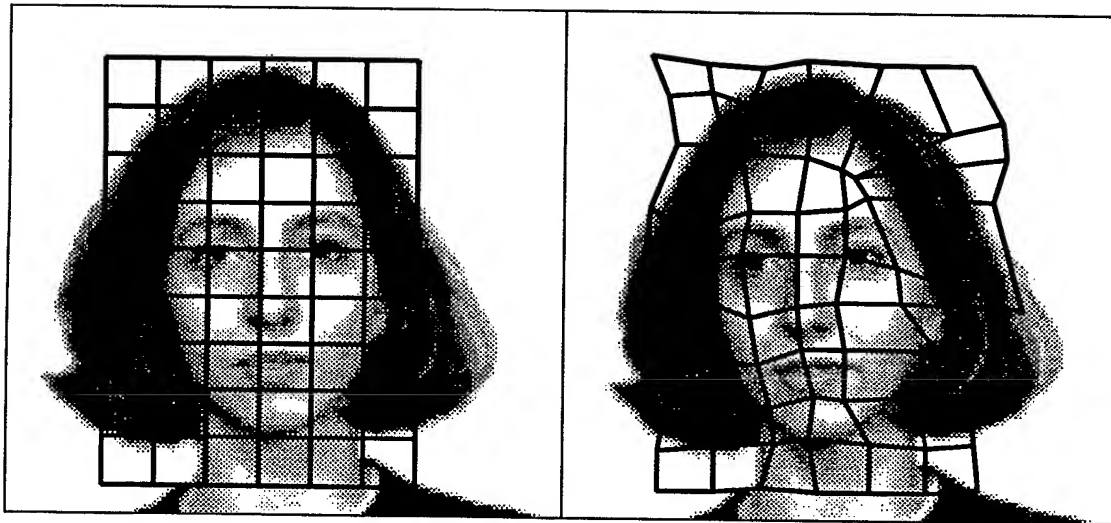


Abbildung 1: (a) Regulärer, (b) irregulärer Graph, bestehend aus jeweils  $7 \times 10$  Knoten.



Abbildung 2: Optimierung des gespeicherten Graphen auf ein kleineres Objekt: (a) gespeichertes Objektbild  $B$  mit gespeichertem Graphen; (b) vorgegebenes Objektbild  $O$  mit gefundenem Graphen nach Optimierung, Phase 1, (c) nach Optimierung, Phase 1+2.